Probabilistic Contextuality in EPR/Bohm-type Systems with Signaling Allowed

Janne V. Kujala
University of Jyväskylä

Ehtibar N. Dzhafarov
Purdue University

Abstract

In this chapter, we review a principled way of defining and measuring contextuality in systems with deterministic inputs and random outputs, recently proposed and developed in (Kujala *et al.*, 2015; Dzhafarov *et al.*, 2015). We illustrate it on systems with two binary inputs and two binary random outputs, the prominent example being the system of two entangled spin-half particles with each particle's spins (random outputs) being measured along one of two directions (inputs). It is traditional to say that such a system exhibits contextuality when it violates Bell-type inequalities. Derivations of Bell-type inequalities, however, are based on the assumption of no-signaling, more generally referred to as marginal selectivity: the distributions of outputs (spins) in Alice's particle do not depend on the inputs (directions chosen) for Bob's particle. In many applications this assumption is not satisfied, so that instead of contextuality one has to speak of direct cross-influences, e.g., of Bob's settings on Alice's spin distributions. While in quantum physics direct cross-influences can sometimes be prevented (e.g., by space-like separation of the two particles), in other applications, especially in behavioral and social systems, marginal selectivity almost never holds. It is unsatisfying that the highly meaningful notion of contextuality is made inapplicable by even slightest violations of marginal selectivity. The new approach rectifies this: it allows one to define and measure contextuality on top of direct cross-influences, irrespective of whether marginal selectivity (no-signaling condition) holds. For systems with two binary inputs and two binary random outputs, contextuality means violation of the classical CHSH inequalities in which the upper bound 2 is replaced with $2(1 + \Delta_0)$, where Δ_0 is a measure of deviation from marginal selectivity.

1 Introduction

In the foundations of quantum physics the notion of contextuality can be formulated in purely probabilistic terms within the framework of the Kolmogorovian probability theory (Larsson, 2002; Khrennikov, 2008a,b; Dzhafarov and

Kujala, 2013a, 2014e,b,c, 2015, 2014a; Dzhafarov, Kujala and Larsson, 2015). The notion applies to any system of random variables recorded under different (mutually incompatible) conditions. Contextuality means that these random variables cannot be "sewn together" into a single system of jointly distributed random variables if one assumes that all or some of them preserve their identity across different conditions. Within the Kolmogorovian framework the existence of this single joint distribution is equivalent to the presentability of all random variables involved as functions of one and the same ("hidden") random variable (Suppes and Zanotti, 1981; Fine, 1982; Dzhafarov and Kujala, 2010).

In spite of its long history (dating from Specker's (1960) example with three boxes, contextuality does not have a standard definition (Kochen and Specker, 1967; Laudisa, 1997; Spekkens, 2008; Kirchmair *et al.*, 2009; Badziąg *et al.*, 2009; Khrennikov, 2009; Cabello, 2013), and is often confounded with such notions as nonlocality and lack of realism (the notions we will not get into in this chapter). All authors who use this term in quantum theory, however, agree on the possibility of detecting contextuality in the spins of entangled particles by violations of Bell-type inequalities (Fine, 1982; Bell, 1964; Clauser *et al.*, 1969). Many other tests have been developed for systems of random variables in and outside quantum physics, notably in psychology (Kujala and Dzhafarov, 2008; Dzhafarov and Kujala, 2012b,a, 2013b). All of these tests are necessary (sometimes also sufficient) conditions for non-contextuality, because of which all of them presuppose or are directly making use of the condition known in psychology as marginal selectivity (Townsend and Schweickert, 1989; Dzhafarov, 2003) and in quantum physics as no-signaling (Cereceda, 2000; Masanes *et al.*, 2006; Oas *et al.*, 2014). In this chapter we use the first term, as more general and purely probabilistic (see Section 7).[1] If marginal selectivity is violated, no "sewing together" of the kind mentioned above is possible.

The problem associated with this fact is that in some cases (including all cases known to us in psychology) violations of marginal selectivity can be readily attributed to the lack of selectivity in the dependence of random variables on various components of the conditions under which they are recorded. If a person is asked to judge brightness and size of a visually presented object, it is not difficult to construct a model in which the judgment of brightness is directly influenced by physical intensity and also directly influenced by object's physical size. In the EPR/Bohm paradigm, if the two measurements of spins in entangled particles are separated by a time-like interval, the spatial axis chosen by Bob (for one of the particles) can in principle initiate a process that will directly influence the spin recorded by Alice (for another particle). We will refer to the dependence of an output distribution on the "wrong" input as a direct cross-influence. The Bell-type inequalities (e.g., in the CHSH form, Clauser *et al.*, 1969) cannot be derived under direct cross-influences, and whether or not they are violated therefore becomes irrelevant.

It seems strange and intellectually unsatisfying, however, that we can detect

---

[1]Within our most recent publications developing the theory, this property is also referred to by the technical name *consistent connectedness*.

contextuality when marginal selectivity holds precisely, but we cannot speak of contextuality at all when it is violated, however slightly. In this chapter we review (in the context of systems with binary inputs and binary random variables as outputs) a recently proposed definition and measure of contextuality (Kujala *et al.*, 2015; Dzhafarov *et al.*, 2015) that overcome this difficulty: even in the presence of direct cross-influences (say, from Bob's setting to Alice's measurements and vice versa) we can detect the presence and compute the degree of contextual influences "on top of" the direct cross-influences. The theory can be generalized to arbitrary systems with deterministic inputs and random outputs, but we do not attempt to present it here. We have made an effort to keep the presentation on a very nontechnical level. This level would be difficult to maintain in a more systematical or more general presentation.

2 The System (α, β, A, B)

Consider a system with two binary inputs, α, β, and two outputs that are binary random variables, A, B. Alice chooses the value of α to be either α_1 or α_2, and she records the corresponding value of A as either $+1$ or -1. Bob chooses the value of β to be either β_1 or β_2, and he records the value of B as either $+1$ or -1. Alice and Bob do this repeatedly in successive trials, so that each input choice and output recording by Alice is paired with an input choice and output recording by Bob. They send their paired choices of inputs and recordings of the outputs to Charlie, who creates four tables of joint distributions: for every $i \in \{1, 2\}$ and $j \in \{1, 2\}$, the distribution is

$$
\begin{array}{|c|cc|c|}
\hline
\phi = (\alpha_i, \beta_j) & B_{ij} = +1 & B_{ij} = -1 & \\
\hline
A_{ij} = +1 & \Pr[A_{ij} = 1, B_{ij} = 1] & \ldots & \Pr[A_{ij} = 1] \\
A_{ij} = -1 & \ldots & \ldots & \ldots \\
\hline
& \Pr[B_{ij} = 1] & \ldots & \\
\hline
\end{array}
\tag{1}
$$

Charlie knows that the only variables that can possibly influence A are α and β, so he labels A recorded under conditions $\phi = (\alpha_i, \beta_j)$ as A_{ij}, allowing thereby A_{ij} to have up to four different distributions. Each of these distributions can be represented by $\Pr[A_{ij} = 1]$, or equivalently by the expected value $\langle A_{ij} \rangle = 2\Pr[A_{ij} = 1] - 1$. The notation B_{ij} for Bob, and the values $\Pr[B_{ij} = 1]$ and $\langle B_{ij} \rangle$ are analogous.

Charlie thus deals with eight random variables,

$$
A_{11}, B_{11}, A_{12}, B_{12}, A_{21}, B_{21}, A_{22}, B_{22}.
\tag{2}
$$

With respect to the joint distribution of A_{ij} and B_{ij}, their individual distributions are referred to as marginal. The joint distribution for (A_{ij}, B_{ij}) is uniquely determined by the two marginal probabilities and the joint probability $\Pr[A_{ij} = +1 \text{ and } B_{ij} = +1]$. Equivalently, it is determined by the two expected values $\langle A_{ij} \rangle, \langle B_{ij} \rangle$ and the product expected value

$$
\langle A_{ij}B_{ij} \rangle = \Pr[A_{ij} = B_{ij}] - \Pr[A_{ij} \neq B_{ij}].
\tag{3}
$$

3 Selectivity of influences and marginal selectivity

Let us assume that Charlie, based on some theory, expects that the dependence of A, B on α, β is selective: Bob's choice of β value does not influence Alice's A and vice versa:

$$
\begin{array}{cc}
\alpha & \beta \\
\downarrow & \downarrow \\
A & B
\end{array}
\tag{4}
$$

This means that A_{i1} and A_{i2} are one and the same random variable for every $i \in \{1, 2\}$, and so are B_{1j} and B_{2j} for every $j \in \{1, 2\}$. Charlie can therefore relabel A_{ij} into A_i and B_{ij} into B_j. But he can also approach this in a more cautious way. He can retain the double indexation and ask the following question: given the eight random variables in (2) of which we know the expectations

$$
(\langle A_{ij} B_{ij} \rangle, \langle A_{ij} \rangle, \langle B_{ij} \rangle), \ i, j \in \{1, 2\},
\tag{5}
$$

can we impose a joint distribution on these eight random variables[2] such that

$$
\begin{array}{ll}
\Pr[A_{i1} \neq A_{i2}] = 0 & \text{for } i \in \{1, 2\} \\
\Pr[B_{1j} \neq B_{2j}] = 0 & \text{for } j \in \{1, 2\}
\end{array} \ ?
\tag{6}
$$

If the answer is affirmative, then the situation is equivalent to the existence of a joint distribution of the single-indexed A_1, B_1, A_2, B_2 such that

$$
(\langle A_i B_j \rangle, \langle A_i \rangle, \langle B_j \rangle) = (\langle A_{ij} B_{ij} \rangle, \langle A_{ij} \rangle, \langle B_{ij} \rangle), \ i, j \in \{1, 2\}.
\tag{7}
$$

However, and this is the reason we call Charlie's approach cautious, the answer does not have to be affirmative. One situation that precludes this is if the following equalities are violated at least for one i or one j:

$$
\langle A_{i1} \rangle = \langle A_{i2} \rangle, \ \langle B_{1j} \rangle = \langle B_{2j} \rangle.
\tag{8}
$$

These equalities represent marginal selectivity of A with respect to changes in β and of B with respect to changes in α. This marginal selectivity is an obvious consequence of (6). If, e.g., $\langle A_{11} \rangle$ were different from $\langle A_{12} \rangle$, then, as Bob changes the value of β from β_1 to β_2, Alice's distribution of A for one and the same choice of $\alpha = \alpha_1$ changes. A_{11} and A_{12} cannot therefore be always equal, contravening (6).

---

[2] To impose a joint distribution on (2) means to create a vector of jointly distributed A'_{11}, B'_{11}, ..., A'_{22}, B'_{22} called a coupling for (2), such that the pairs $\left(A'_{ij}, B'_{ij}\right)$ have the same distributions as (A_{ij}, B_{ij}) for all $i, j \in \{1, 2\}$. No other subset of (2) has a joint distribution. In this chapter we conveniently confuse random variables and their primed counterparts. See (Dzhafarov and Kujala, 2013a, 2014e,b,c, 2015, 2010, forthcoming) for detailed discussions.

In situations like this Charlie is forced then to revise his model (4) in favor of

$$\begin{array}{ccc} \alpha & & \beta \\ \big\downarrow & \times & \big\downarrow \\ A & & B \end{array} \tag{9}$$

This can be referred to as a model with direct cross-influences: the distribution (hence also identity) of the outputs is allows to be influenced by "wrong" inputs ("wrong" from the point of view of the Charlie's original theory[3]).

4 Contextuality under marginal selectivity

There is also another possibility for Charlie's question to have a negative answer. The marginal selectivity requirement may very well be satisfied, but the observed expectations (5) may be incompatible with the hypothesis (6). The incompatibility means that a joint distribution of the eight random variables (2) that accords with both (5) and (6) does not exist. This understanding of contextuality was first utilized by Larsson (2002). It helps to understand the essence of all Bell-type theorems. Stated in the form convenient for our purposes, the theorem that applies to all systems with two binary inputs and two binary random outputs Fine (1982) says:

Theorem 4.1 (Fine, 1982). *The observed expectations (5) are compatible with the identity connections (6) if and only if marginal selectivity (8) is satisfied for all $i, j \in \{1, 2\}$, and*

$$\max_{i,j \in \{1,2\}} |\langle A_{11}B_{11} \rangle + \langle A_{12}B_{12} \rangle + \langle A_{21}B_{21} \rangle + \langle A_{22}B_{22} \rangle - 2 \langle A_{ij}B_{ij} \rangle| \leq 2. \tag{10}$$

The term "connections" used in this formulation (Dzhafarov and Kujala, 2013a, 2014e,b,c) refers to the unobservable pairs

$$(A_{11}, A_{12}), (A_{21}, A_{22}), (B_{11}, B_{21}), (B_{12}, B_{22}). \tag{11}$$

---

[3]This is the "subjective", or theory-laden aspect of the notion of contextuality: this notion acquires its meaning only in relation to some model, in this case represented by (4), that describes the system the way it "ought to be" or predicted to be by some theory. We will not elaborate, but this accords with our view (Dzhafarov and Kujala, 2014a) that while probabilities are objective, the identities of random variables are theory-laden.

Their unobservable joint distributions are given by

	$A_{i2} = +1$	$A_{i2} = -1$	
$A_{i1} = +1$	$\Pr[A_{i1} = 1, A_{i2} = 1]$	...	$\Pr[A_{i1} = 1]$
$A_{i1} = -1$	...	...	...
	$\Pr[A_{i2} = 1]$	...	

	$B_{2j} = +1$	$B_{2j} = -1$	
$B_{1j} = +1$	$\Pr[B_{1j} = 1, B_{2j} = 1]$	...	$\Pr[B_{1j} = 1]$
$B_{1j} = -1$	...	...	...
	$\Pr[B_{2j} = 1]$	...	

$$(12)$$

for $i, j \in \{1, 2\}$. If (6) holds, i.e., the entries on the minor diagonals of the tables are zero, then the connections are called the identity ones.

The compatibility of connections with the observed expectations (uniquely defining the observed distributions) means that each of the 2^8 possible combinations

$$A_{11} = \pm 1, B_{11} = \pm 1, \ldots, A_{22} = \pm 1, B_{22} = \pm 1$$

is assigned a probability, so that the probabilities for all combinations containing, say, $A_{12} = 1$ and $B_{12} = -1$ sum to the observed $\Pr[A_{12} = 1, B_{12} = -1]$; and the probabilities for all combinations containing, say, $B_{12} = 1$ and $B_{22} = 1$ equals the hypothetical (unobservable) connection probability $\Pr[B_{12} = 1, B_{22} = 1]$.

The inequalities (10), in physics referred to as CHSH inequalities, can be violated, and they are de facto violated if A and B are spins of two entangled particles under certain choices of spatial axes (α and β) along which they are measured (Aspect *et al.*, 1981, 1982; Weihs *et al.*, 1998). When these inequalities are violated while marginal selectivity is satisfied, we speak of contextuality: Alice's output A under her choice of α_1 does not change its distribution depending on Bob's choice of β_1 or β_2, but A_{11} and A_{12} still cannot be considered one and the same random variable (it should not come as a surprise that different random variables can have the same distribution).

In the diagram below the interrupted lines indicate contextual influences: the dependence of identities of identically distributed random variables on the "wrong" inputs:

$$(13)$$

When the inequalities (10) are violated, a measure of contextuality can be easily designed as follows. If (6) were compatible with the observed expectations (5), then (by definition) Charlie could construct a joint distribution of the random variables (2) in which

$$\Delta = \Pr[A_{11} \neq A_{12}] + \Pr[A_{21} \neq A_{22}] + \Pr[B_{11} \neq B_{21}] + \Pr[B_{12} \neq B_{22}] \quad (14)$$

equals zero. If (6) is incompatible with (5), then this Δ cannot be zero in any joint distribution imposed on (2). It is natural therefore to adopt the following

Definition 4.2. Under marginal selectivity, the degree of contextuality in a system with given observed expectations (5) is the minimal value of Δ in (14) for which a joint distribution for (2) exists.

As it turns out, this minimal value of Δ equals

$$\Delta_{\min} = \max\{0, \Delta_{\text{CHSH}}\}, \tag{15}$$

where

$$\Delta_{\text{CHSH}} = \\ \tfrac{1}{2}\max_{i,j\in\{1,2\}} |\langle A_{11}B_{11}\rangle + \langle A_{12}B_{12}\rangle + \langle A_{21}B_{21}\rangle + \langle A_{22}B_{22}\rangle - 2\langle A_{ij}B_{ij}\rangle| - 1 \tag{16}$$

is ($1/2$ times) the violation of the CHSH inequalities. This is a special case of the formula derived later in Theorem 6.1 without the assumption of marginal selectivity.

As an example, let the observed expectations be at the Tsirelson bounds (Tsirelson, 1980; Landau, 1987). Then $\Delta_{\min}$ is $\sqrt{2} - 1$. The largest possible value of $\Delta_{\min}$ is 1.

5 Contextuality on top of direct cross-influences

The definition of contextuality given above does not work for the situation depicted in (9), where marginal selectivity is not satisfied. In this case we have direct cross-influences from "wrong" inputs, and this precludes the possibility that Δ in (14) is zero. In fact, we have the simple

Theorem 5.1. *Given the observed expectations* $(\langle A_{ij}\rangle, \langle B_{ij}\rangle)_{i,j\in\{1,2\}}$, *the minimum possible value for* Δ *in (14) is*

$$\begin{aligned} \Delta_0 = \tfrac{1}{2}(\, &|\langle A_{11}\rangle - \langle A_{12}\rangle| + |\langle A_{21}\rangle - \langle A_{22}\rangle| + \\ &|\langle B_{11}\rangle - \langle B_{21}\rangle| + |\langle B_{12}\rangle - \langle B_{22}\rangle|). \end{aligned} \tag{17}$$

Proof. We minimize Δ if we minimize separately $\Pr[A_{11} \neq A_{12}]$, $\Pr[A_{21} \neq A_{22}]$, $\Pr[B_{11} \neq B_{21}]$, and $\Pr[B_{12} \neq B_{22}]$. Consider, e.g., the distribution of the connection (A_{11}, A_{12}):

	$A_{12} = +1$	$A_{12} = -1$
$A_{11} = +1$	$\Pr[A_{11} = 1, A_{12} = 1]$	$\Pr[A_{11} = 1] - \Pr[A_{11} = 1, A_{12} = 1]$
$A_{11} = -1$	$\Pr[A_{12} = 1] - \Pr[A_{11} = 1, A_{12} = 1]$	$\ldots$

$$\tag{18}$$

The largest possible value for the probability $\Pr[A_{11} = 1, A_{12} = 1]$ is

$$\min\{\Pr[A_{11} = 1], \Pr[A_{12} = 1]\},$$

whence the minimum of $\Pr[A_{11} \neq A_{12}]$, which is the sum of the entries on the minor diagonal, is $|\Pr[A_{11} = 1] - \Pr[A_{12} = 1]| = \tfrac{1}{2}|\langle A_{11}\rangle - \langle A_{12}\rangle|$. $\qquad\square$

Under the marginal selectivity we have $\Delta_0 = 0$, and we speak of contextuality if the minimal value of Δ that is compatible with the observed expectations (5) is greater than $\Delta_0 = 0$. In the general case $\Delta_0 > 0$, and we need a more general definition of contextuality. The idea is simple. If $\Delta_0 > 0$, we have direct cross-influences (9), and if $\Delta = \Delta_0$ is compatible with the observed expectations (5), then no contextuality is involved: direct cross-influences is all one needs to account for the system's behavior. If however $\Delta = \Delta_0$ is not compatible with the observed expectations (5), then we can speak of contextuality "on top of" the direct cross-influences. The natural measure of the degree of contextuality then is given by

Definition 5.2. The degree of contextuality in a system with given observed expectations (5) is $\Delta_{\min} - \Delta_0$, where $\Delta_{\min}$ is the minimal value of Δ in (14) for which a joint distribution for (2) exists.

6 General formula for contextuality

We now need to derive a formula for $\Delta_{\min}$ of which (15) is a special case.

Theorem 6.1. *The minimum possible value $\Delta_{\min}$ for Δ that is compatible with the observed expectations (5) is*

$$\Delta_{\min} = \max\{\Delta_0, \Delta_{\mathrm{CHSH}}\}, \tag{19}$$

where Δ_0 is given in (17) and Δ_{CHSH} in (16).

Proof. By Lemma 8.2 (a computer-assisted result detailed in the next section), Δ is compatible with the observed $(\langle A_{ij}B_{ij}\rangle, \langle A_{ij}\rangle, \langle B_{ij}\rangle)_{i,j\in\{1,2\}}$ if and only if it satisfies

$$\Delta \geq -1 + \tfrac{1}{2}s_1\left(\langle A_{11}B_{11}\rangle, \langle A_{12}B_{12}\rangle, \langle A_{21}B_{21}\rangle, \langle A_{22}B_{22}\rangle\right), \tag{20}$$

$$\Delta \geq \tfrac{1}{2}(|\langle A_{11}\rangle - \langle A_{12}\rangle| + |\langle A_{21}\rangle - \langle A_{22}\rangle| + \\ |\langle B_{11}\rangle - \langle B_{21}\rangle| + |\langle B_{12}\rangle - \langle B_{22}\rangle|), \tag{21}$$

$$\Delta \leq 4 - \left[-1 + \tfrac{1}{2}s_1\left(\langle A_{11}B_{11}\rangle, \langle A_{12}B_{12}\rangle, \langle A_{21}B_{21}\rangle, \langle A_{22}B_{22}\rangle\right)\right], \tag{22}$$

$$\Delta \leq 4 - \tfrac{1}{2}(|\langle A_{11}\rangle + \langle A_{12}\rangle| + |\langle A_{21}\rangle + \langle A_{22}\rangle| + \\ |\langle B_{11}\rangle + \langle B_{21}\rangle| + |\langle B_{12}\rangle + \langle B_{22}\rangle|), \tag{23}$$

where $s_1(\cdots)$ is defined in (39) in Sec. 8 below and is equal to the max $|\ldots|$-part of (16). These inequalities are always mutually compatible, whence $\Delta_{\min}$ is the larger of the two right-hand expressions in (20) and (21). $\qquad\square$

It follows that $\Delta_{\min} - \Delta_0$ is always nonnegative, and Definition 5.2 is well-constructed: $\Delta_{\min} - \Delta_0 = 0$ indicates no contextuality, $\Delta_{\min} - \Delta_0 > 0$ indicates contextuality on top of the direct cross-influences.

We can present the notion of (non-)contextuality in as close a form as possible to the traditional CHSH inequalities:

8

Theorem 6.2. *The system exhibits no contextuality if and only if*

$$|\langle A_{11}B_{11}\rangle + \langle A_{12}B_{12}\rangle + \langle A_{21}B_{21}\rangle - \langle A_{22}B_{22}\rangle| \leq 2\left(1+\Delta_0\right),$$
$$|\langle A_{11}B_{11}\rangle + \langle A_{12}B_{12}\rangle - \langle A_{21}B_{21}\rangle + \langle A_{22}B_{22}\rangle| \leq 2\left(1+\Delta_0\right),$$
$$|\langle A_{11}B_{11}\rangle - \langle A_{12}B_{12}\rangle + \langle A_{21}B_{21}\rangle + \langle A_{22}B_{22}\rangle| \leq 2\left(1+\Delta_0\right),$$
$$|-\langle A_{11}B_{11}\rangle + \langle A_{12}B_{12}\rangle + \langle A_{21}B_{21}\rangle + \langle A_{22}B_{22}\rangle| \leq 2\left(1+\Delta_0\right),$$

$$(24)$$

where Δ_0 is the natural measure of violation of marginal selectivity, (17). If at least one of these inequalities is violated, then the largest difference between the left-hand side and $2\left(1+\Delta_0\right)$ is the degree of contextuality (after scaling by $1/2$).

The maximum value attainable by one of the linear combinations in (24) is 4. It follows that the system exhibits no contextuality if the violation of marginal selectivity Δ_0 in it is not less than 1. Put differently, if $\Delta_0 \geq 1$, any observed distributions of random variables can be accounted for in terms of direct cross-influences, with no contextuality involved.

7 Consequences of the new definition of contextuality

The notion of contextuality was presented in Introduction to mean that random variables recorded under mutually incompatible conditions cannot be "sewn together" into a single system of jointly distributed random variables, provided one assumes that all or some of them preserve their identity across different conditions. We should now relax the assumption clause:

> contextuality means that random variables recorded under mutually incompatible conditions cannot be "sewn together" into a single system of jointly distributed random variables, provided one assumes that their identity across different conditions changes as little as possibly allowed by direct cross-influences (equivalently, by observed deviations from marginal selectivity).

As mentioned in Introduction, marginal selectivity is rarely satisfied outside quantum physics, and, in particular, is almost always violated in psychological experiments. Consider, e.g., a double-detection experiment, where a participant is presented two side-by-side flashes of light (left and right) and asked to say "Yes/No" to the question "Is there a flash on the left?" and another "Yes/No" to the question "Is there a flash on the right?". Each flash can be presented at two intensity levels: zero (no flash) and some very small value $s > 0$. We have therefore four conditions: $(0,0), (0,s), (s,0), (s,s)$. Denoting the response about the left stimulus by A and he response about the right stimulus by B, we get the eight random variables $A_{00}, B_{00}, \ldots, A_{ss}, B_{ss}$. The situation is formally identical to the Alice-Bob paradigm. The "normative" diagram (4), with α, β being the two flash intensities, is very likely to be violated on the level of marginal probabilities: the answer about the left flash will almost certainly be

influenced by the intensity of the right flash, and vice versa. Our definition of contextuality, however, allows one to determine whether contextuality is there on top of these direct cross-influences.

Another example is taken from the work by Aerts *et al.* (2013). They estimated the probabilities with which people chose one of two presented to them animal names and one of two presented to them animal sounds. The results were as follows:

Probability estimates from Table 1 of (Aerts *et al.*, 2013). [†]

$\phi = (\alpha_1, \beta_1)$	$B_{11} =$ Growls	$B_{11} =$ Whinnies		$\phi = (\alpha_1, \beta_2)$	$B_{12} =$ Snorts	$B_{12} =$ Meows	
$A_{11} =$ Horse	.049	.630	.679	$A_{12} =$ Horse	.593	.025	.618
$A_{11} =$ Bear	.259	.062	.321	$A_{12} =$ Bear	.296	.086	.382
	.308	.692			.889	.111	

$\phi = (\alpha_2, \beta_1)$	$B_{21} =$ Growls	$B_{21} =$ Whinnies		$\phi = (\alpha_2, \beta_2)$	$B_{22} =$ Snorts	$B_{22} =$ Meows	
$A_{21} =$ Tiger	.778	.086	.864	$A_{22} =$ Tiger	.148	.086	.234
$A_{21} =$ Cat	.086	.049	.135	$A_{22} =$ Cat	.099	.667	.766
	.864	.135			.247	.753	

[†] Based on 81 respondents per table.

Here, α indicates one of the two animal dichotomies offered ($\alpha_1 =$ Horse or Bear, $\alpha_2 =$ Tiger or Cat), and β analogously indicates one of two animal sound dichotomies. The value of Δ_{CHSH} given by (16) equals 0.210 here, and Aerts et al. report it as evidence in favor of contextuality (note that the CHSH bound of 2 corresponds to $\Delta_{\mathrm{CHSH}} = 0$). We criticized this conclusion (Dzhafarov and Kujala, 2014d) by pointing out that the derivation of the CHSH inequalities is not valid without marginal selectivity, and the latter is clearly violated in the data: e.g., $\Pr[B_{12} = \mathrm{Snorts}] = 0.889$ while $\Pr[B_{22} = \mathrm{Snorts}] = 0.247$.

We can now amend our criticism: the computation of Δ_{CHSH} is meaningful even if marginal selectivity is contravened. One has, however, to compare Δ_{CHSH} to Δ_0 of (17) rather than to zero, and to compute $\max\{\Delta_0, \Delta_{\mathrm{CHSH}}\} - \Delta_0$ as the measure of contextuality. Unfortunately for the Aerts et al.'s conclusions, Δ_0 in their data is too large (1.889) to allow for nonzero contextuality.

In quantum physics, the no-signaling condition (a special case of marginal selectivity) can be ensured by separating the outputs from the "wrong" inputs by space-like intervals. There are, however, some indications that in the well-known experiments by Weihs *et al.* (1998), where space-like separation is claimed to be the case, some violations of marginal selectivity were observed (Adenier and Khrennikov, 2007). If so, and whatever the physical cause of these violations, our new approach provides a way of testing whether contextuality is still present in the data.

Signaling is natural to assume in Leggett–Garg (1985) -type systems, with three binary random variables X, Y, Z tied to three successive moments of time, $t_1 < t_2 < t_3$. Any two of these three random variables can be measured together, in one experiment, but not all three of them. If X and Z are measured together, then (in accordance with our general approach, see Dzhafarov and Kujala, 2014e,b,c, 2015, 2014a) the identity of X as a random variable may be different from the identity of X when measured together with Y. This means that X in the two situations should be labelled differently, say, X_{13} and X_{12}, respectively (based on the time moments involved). Analogously, we have Y_{12} and Y_{23} depending on whether Y is measured together with X or with Z; and we have Z_{13} and Z_{23}.

Suppes and Zanotti (1981) have shown that given uniform marginals, an equivalent condition for the existence of a joint distribution of

$$X_{12}, X_{13}, Y_{12}, Y_{23}, Z_{13}, Z_{23} \tag{25}$$

under the constraint $X_{12} = X_{13}$, $Y_{12} = Y_{23}$, $Z_{13} = Z_{23}$ is

$$
\begin{aligned}
-1 &\leq \langle X_{12}Y_{12}\rangle + \langle Y_{23}Z_{23}\rangle + \langle X_{13}Z_{13}\rangle \\
&\leq 1 + 2\max\left\{\langle X_{12}Y_{12}\rangle, \langle Y_{23}Z_{23}\rangle, \langle X_{13}Z_{13}\rangle\right\}.
\end{aligned}
\tag{26}
$$

As a side product of our analysis, we show that this inequality in fact holds for arbitrary marginals as well and we generalize the inequalities to the signaling case.

Theorem 7.1. *The minimum possible value $\Delta'_{\min}$ for*

$$\Delta' = \Pr\left[X_{12} \neq X_{13}\right] + \Pr\left[Y_{12} \neq Y_{23}\right] + \Pr\left[Z_{13} \neq Z_{23}\right] \tag{27}$$

that is compatible with the observed expectations

$$\langle X_{12}Y_{12}\rangle, \langle X_{13}Z_{13}\rangle, \langle Y_{23}Z_{23}\rangle, \langle X_{12}\rangle, \langle X_{13}\rangle, \langle Y_{12}\rangle, \langle Y_{23}\rangle, \langle Z_{13}\rangle, \langle Z_{23}\rangle \tag{28}$$

is

$$\Delta'_{\min} = \max\left\{\Delta'_0, \Delta'_{\mathrm{SZ}}\right\}, \tag{29}$$

where

$$\Delta'_0 = \frac{1}{2}\left(|\langle X_{12}\rangle - \langle X_{13}\rangle| + |\langle Y_{12}\rangle - \langle Y_{23}\rangle| + |\langle Z_{13}\rangle - \langle Z_{23}\rangle|\right) \tag{30}$$

is the natural measure of the violation of marginal selectivity and

$$
\Delta'_{\mathrm{SZ}} = -\tfrac{1}{2} + \tfrac{1}{2}\max\left\{
\begin{aligned}
&\langle X_{12}Y_{12}\rangle + \langle X_{13}Z_{13}\rangle - \langle Y_{23}Z_{23}\rangle, \\
&\langle X_{12}Y_{12}\rangle - \langle X_{13}Z_{13}\rangle + \langle Y_{23}Z_{23}\rangle, \\
&-\langle X_{12}Y_{12}\rangle + \langle X_{13}Z_{13}\rangle + \langle Y_{23}Z_{23}\rangle, \\
&-\langle X_{12}Y_{12}\rangle - \langle X_{13}Z_{13}\rangle - \langle Y_{23}Z_{23}\rangle
\end{aligned}
\right\}
\tag{31}
$$

is ($^1\!/_2$ times) the maximum violation of the Suppes–Zanotti inequalities (26).

Proof. By Lemma 8.5 of the next section, Δ' is compatible with the observed expectations (28) if and only if it satisfies

$$\Delta' \geq -\tfrac{1}{2} + \tfrac{1}{2} s_1 \left(\langle X_{12} Y_{12} \rangle, \langle Y_{23} Z_{23} \rangle, \langle X_{13} Z_{13} \rangle \right), \tag{32}$$

$$\Delta' \geq \tfrac{1}{2} \left(|\langle X_{12} \rangle - \langle X_{13} \rangle| + |\langle Y_{12} \rangle - \langle Y_{23} \rangle| + |\langle Z_{13} \rangle - \langle Z_{23} \rangle| \right), \tag{33}$$

$$\Delta' \leq 3 - \left[-\tfrac{1}{2} - \tfrac{1}{2} s_1 \left(\langle X_{12} Y_{12} \rangle, \langle Y_{23} Z_{23} \rangle, \langle X_{13} Z_{13} \rangle \right) \right], \tag{34}$$

$$\Delta' \leq 3 - \tfrac{1}{2} \left(|\langle X_{12} \rangle + \langle X_{13} \rangle| + |\langle Y_{12} \rangle + \langle Y_{23} \rangle| + |\langle Z_{13} \rangle + \langle Z_{23} \rangle| \right). \tag{35}$$

These inequalities are always mutually compatible, whence $\Delta'_{\min}$ is the larger of the two right-hand expressions in (32) and (33). $\square$

Definition 7.2. The degree of contextuality in a system with given observed expectations (28) is $\Delta'_{\min} - \Delta'_0$, where $\Delta'_{\min}$ is the minimal value of Δ' in (27) for which a joint distribution for (25) exists.

Using essentially the same reasoning as for the EPR/Bohm paradigm, we come to the following

Theorem 7.3. *A Leggett–Garg-type systems exhibits no contextuality if and only if*

$$\begin{aligned}
\langle X_{12} Y_{12} \rangle + \langle Y_{23} Z_{23} \rangle - \langle X_{13} Z_{13} \rangle &\leq 1 + 2\Delta'_0, \\
\langle X_{12} Y_{12} \rangle - \langle Y_{23} Z_{23} \rangle + \langle X_{13} Z_{13} \rangle &\leq 1 + 2\Delta'_0, \\
- \langle X_{12} Y_{12} \rangle + \langle Y_{23} Z_{23} \rangle + \langle X_{13} Z_{13} \rangle &\leq 1 + 2\Delta'_0, \\
- \langle X_{12} Y_{12} \rangle - \langle Y_{23} Z_{23} \rangle - \langle X_{13} Z_{13} \rangle &\leq 1 + 2\Delta'_0.
\end{aligned} \tag{36}$$

The largest in absolute value breach of one of these boundaries then can be taken as a measure of contextuality.

Inequalities (36) can also be equivalently rewritten closer to the Suppes–Zanotti (1981) formulation:

$$\begin{aligned}
-1 - 2\Delta'_0 &\leq \langle X_{12} Y_{12} \rangle + \langle Y_{23} Z_{23} \rangle + \langle X_{13} Z_{13} \rangle \\
&\leq 1 + 2\Delta'_0 + 2 \max \left\{ \langle X_{12} Y_{12} \rangle, \langle Y_{23} Z_{23} \rangle, \langle X_{13} Z_{13} \rangle \right\}.
\end{aligned} \tag{37}$$

8 Technical details

In this section, we give the technical details of the computer-assisted results used above. Refer to Fig. 1 for a graphical representation of the connections and observed pairs of random variables in the system.[4]

---

[4]Based on our most recent theoretical results (Kujala *et al.*, 2015; Kujala and Dzhafarov, 2015), the computer-assisted proofs for the systems considered here can in fact be obtained analytically as well. However, the principles of computer-assisted proof laid out here are applicable in systems that are not covered by the analytical results.

Lemma 8.1. *The necessary and sufficient condition for the connection expectations* $(\langle A_{i1}A_{i2}\rangle, \langle B_{1j}B_{2j}\rangle)_{i,j\in\{1,2\}}$ *to be compatible with the observed expectations*

$$(\langle A_{ij}B_{ij}\rangle, \langle A_{ij}\rangle, \langle B_{ij}\rangle)_{i,j\in\{1,2\}}$$

is

$$
\begin{aligned}
&s_0\left(\langle A_{11}B_{11}\rangle, \langle A_{12}B_{12}\rangle, \langle A_{21}B_{21}\rangle, \langle A_{22}B_{22}\rangle\right)\\
&\quad \leq 6 - s_1\left(\langle A_{11}A_{12}\rangle, \langle B_{11}B_{21}\rangle, \langle A_{21}A_{22}\rangle, \langle B_{12}B_{22}\rangle\right),\\
&s_1\left(\langle A_{11}B_{11}\rangle, \langle A_{12}B_{12}\rangle, \langle A_{21}B_{21}\rangle, \langle A_{22}B_{22}\rangle\right)\\
&\quad \leq 6 - s_0\left(\langle A_{11}A_{12}\rangle, \langle B_{11}B_{21}\rangle, \langle A_{21}A_{22}\rangle, \langle B_{12}B_{22}\rangle\right),
\end{aligned}
\tag{38}
$$

where

$$
\begin{aligned}
s_0\left(a,b,c,d\right) &= \max\left\{(\pm a \pm b \pm c \pm d): \text{ the number of minuses is even}\right\},\\
s_1\left(a,b,c,d\right) &= \max\left\{(\pm a \pm b \pm c \pm d): \text{ the number of minuses is odd}\right\}.
\end{aligned}
\tag{39}
$$

Proof. The joint distribution of the eight random variables

$$A_{11}, B_{11}, A_{12}, B_{12}, A_{21}, B_{21}, A_{22}, B_{22}$$

is fully described by the vector $\mathbf{q} \in [0,1]^n$, $q_1 + \cdots + q_n = 1$, consisting of the probabilities of the $n = 2^8 = 256$ different combinations of the values of the eight random variables. We then define a vector $\mathbf{p} \in [0,1]^m$, $m = 32$, consisting of the 16 observable probabilities $\Pr[A_{ij} = a, \ B_{ij} = b]$ for $a, b \in \{-1,1\}$, $i, j \in \{1,2\}$ and the 16 connection probabilities given by $\Pr[A_{i1} = a, \ A_{i2} = a']$ and $\Pr[B_{1j} = b, \ B_{2j} = b']$ for $a, a', b, b' \in \{-1,1\}$ and $i, j \in \{1,2\}$. As every element of $\mathbf{p}$ is a (2-)marginal probability of the joint represented by $\mathbf{q}$, there exists a binary marix $M \in \{0,1\}^{m\times n}$ such that

$$\mathbf{p} = M\mathbf{q}. \tag{40}$$

It follows that the observable probabilities $p_1, \ldots, p_{16}$ are compatible with the connection probabilities $p_{17}, \ldots, p_{32}$ if and only if there exists an n-vector $\mathbf{q} \geq 0$ such that (40) holds. As described in (Dzhafarov and Kujala, 2013a, Text S3), the set of vectors $\mathbf{p}$ satisfying this constraint forms a polytope whose vertices are given by the columns of M and whose half space representation can be obtained by a facet enumeration algorithm. As also described in (Dzhafarov and Kujala, 2013a), this halfspace representation consists of 160 inequalities and 16 equations in $p_1, \ldots, p_{32}$. The 16 equations correspond to the requirement that the 1-marginals of the observable probabilities agree with those of the connections and that the observable probabilities are properly normalized.

Expressing the probabilities in the vector $\mathbf{p}$ in terms of the observable and connection expectations $(\langle A_{ij}B_{ij}\rangle, \langle A_{ij}\rangle, \langle B_{ij}\rangle, \langle A_{i1}A_{i2}\rangle, \langle B_{1j}B_{2j}\rangle)$, $i, j \in \{1,2\}$, the 16 equations become identically true (the parameterization already guarantees them), and of the 160 inequalities, 128 turn into exactly those represented by (38) and the remaining 32 are trivial constraints of the form

$$-1 + |\langle A\rangle + \langle B\rangle| \leq \langle AB\rangle \leq 1 - |\langle A\rangle - \langle B\rangle| \tag{41}$$

for the 8 pairs of random variables involved in (38). The trivial constraints correspond to the implicit requirement that the observable and connection probabilities are nonnegative and thus they need not be explicitly shown in the statement of the theorem. $\square$

This proof is different from the similar result in (Dzhafarov and Kujala, 2013a) in that the parameterization for the probabilities in $\mathbf{p}$ is more general (allowing for arbitrary marginals of the eight random variables) and so we obtain a more general condition for the compatibility of observable and connection probabilities than before. It should be noted that although the expectations $\langle A_{ij} \rangle, \langle B_{ij} \rangle, i,j \in \{1,2\}$ do not explicitly appear in (38), they are still present in the 32 implicit constraints.

Lemma 8.2. *If the connection expectations* $(\langle A_{i1}A_{i2} \rangle, \langle B_{1j}B_{2j} \rangle)_{i,j \in \{1,2\}}$ *are compatible with the observed expectations* $(\langle A_{ij}B_{ij} \rangle, \langle A_{ij} \rangle, \langle B_{ij} \rangle)_{i,j \in \{1,2\}}$, *then, with Δ defined as in (14),*

$$
\begin{aligned}
\Delta &\geq -1 + \tfrac{1}{2} s_1 \left(\langle A_{11}B_{11} \rangle, \langle A_{12}B_{12} \rangle, \langle A_{21}B_{21} \rangle, \langle A_{22}B_{22} \rangle \right), \\
\Delta &\geq \tfrac{1}{2}(|\langle A_{11} \rangle - \langle A_{12} \rangle| + |\langle A_{21} \rangle - \langle A_{22} \rangle| + \\
&\qquad |\langle B_{11} \rangle - \langle B_{21} \rangle| + |\langle B_{12} \rangle - \langle B_{22} \rangle|), \\
\Delta &\leq 4 - \left[-1 + \tfrac{1}{2} s_1 \left(\langle A_{11}B_{11} \rangle, \langle A_{12}B_{12} \rangle, \langle A_{21}B_{21} \rangle, \langle A_{22}B_{22} \rangle \right) \right], \\
\Delta &\leq 4 - \tfrac{1}{2}(|\langle A_{11} \rangle + \langle A_{12} \rangle| + |\langle A_{21} \rangle + \langle A_{22} \rangle| + \\
&\qquad |\langle B_{11} \rangle + \langle B_{21} \rangle| + |\langle B_{12} \rangle + \langle B_{22} \rangle|).
\end{aligned}
\tag{42}
$$

Conversely, if these inequalities are satisfied for a given value of Δ, then the connection expectations $(\langle A_{i1}A_{i2} \rangle, \langle B_{1j}B_{2j} \rangle)_{i,j \in \{1,2\}}$ *can always be chosen so that they are compatible with the observable expectations* $(\langle A_{ij}B_{ij} \rangle, \langle A_{ij} \rangle, \langle B_{ij} \rangle)_{i,j \in \{1,2\}}$ *and yield the given value of Δ in (14).*

Proof. Given the 160 inequalities (including the 32 implicit inequalities) of Lemma 8.1 characterizing the compatibility of the connection expectations with the observable expectations, we amend this linear system with the equation (14) defining Δ written in terms of the expectations

$$
\left(\langle A_{i1}A_{i2} \rangle, \langle B_{1j}B_{2j} \rangle, \langle A_{ij} \rangle, \langle B_{ij} \rangle \right)_{i,j \in \{1,2\}}.
$$

Then, we use this equation to eliminate one of the connection expectation variables $(\langle A_{i1}A_{i2} \rangle, \langle B_{1j}B_{2j} \rangle)_{i,j \in \{1,2\}}$ from the system (by solving the variable from the equation and then substituting the solution everywhere else). After that, we eliminate the three remaining connection expectation variables one by one using the Fourier–Motzkin elimination algorithm (see Theorem 8.3 below). After the elimination of each variable, we remove any redundant inequalities from the system by linear programming using the algorithm described in (Dzhafarov and Kujala, 2013a, Text S3). After having eliminated all connection expectation variables, we are left with the system (42) (and implicit constraints of the form
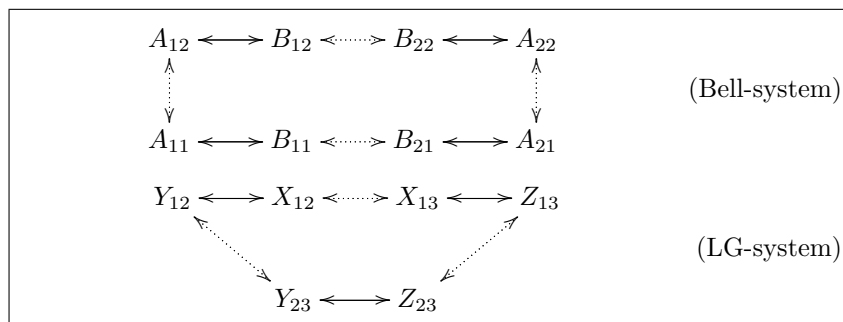
14

$$A_{12} \longleftrightarrow B_{12} \longleftdashrightarrow B_{22} \longleftrightarrow A_{22}$$

(Bell-system)

$$A_{11} \longleftrightarrow B_{11} \longleftdashrightarrow B_{21} \longleftrightarrow A_{21}$$

$$Y_{12} \longleftrightarrow X_{12} \longleftdashrightarrow X_{13} \longleftrightarrow Z_{13}$$

(LG-system)

$$Y_{23} \longleftrightarrow Z_{23}$$

Figure 1: Random variables involved in the Bell-system and LG-system. The pairs of random variables whose joint distributions are empirically observed, e.g., (A_{12}, B_{12}) and (X_{12}, Y_{12}), are indicated by solid double-arrows. The pairs of random variables forming probabilistic connections (with unobservable joint distributions) are indicated by point double-arrows, e.g., (A_{11}, A_{12}) and (X_{12}, X_{13}).

(41) for the pairs (A_{ij}, B_{ij}), $i, j \in \{1, 2\}$). The Fourier–Motzkin elimination algorithm guarantees that the resulting system has a solution precisely when the original system has a solution with *some* values of the eliminated variables. $\square$

Theorem 8.3 (Fourier–Motzkin elimination). *Given a system of linear inequalities in the variables x and $\mathbf{y} = y_1, \ldots, y_n$, the system can always be rearranged in the following form*

$$
\begin{aligned}
x &\geq \mathbf{l}_i \cdot \mathbf{y}, & i &= 1, \ldots, n_{\mathbf{l}}, \\
x &\leq \mathbf{u}_i \cdot \mathbf{y}, & i &= 1, \ldots, n_{\mathbf{u}}, \\
0 &\leq \mathbf{n}_i \cdot \mathbf{y}, & i &= 1, \ldots, n_{\mathbf{n}},
\end{aligned}
\tag{43}
$$

where $\mathbf{l}_1, \ldots, \mathbf{l}_{n_l}, \mathbf{u}_1, \ldots, \mathbf{u}_{n_u}, \mathbf{n}_1, \ldots, \mathbf{n}_{n_n} \in \mathbb{R}^n$. Furthermore, given $\mathbf{y} \in \mathbb{R}$, this system is solved by $\mathbf{y}$ and some $x \in \mathbb{R}$ if and only if the following system is solved by $\mathbf{y}$:

$$
\begin{aligned}
\mathbf{l}_i \cdot \mathbf{y} &\leq \mathbf{u}_j \cdot \mathbf{y}, & i &= 1, \ldots, n_{\mathbf{l}}, \ j = 1, \ldots, n_{\mathbf{u}}, \\
0 &\leq \mathbf{n}_i \cdot \mathbf{y}, & i &= 1, \ldots, n_{\mathbf{n}}.
\end{aligned}
\tag{44}
$$

Lemma 8.4. *The necessary and sufficient condition for the connection expectations $\langle X_{12}X_{13} \rangle$, $\langle Y_{12}Y_{23} \rangle$, $\langle Z_{13}Z_{23} \rangle$ to be compatible with the observed expectations $\langle X_{12}Y_{12} \rangle$, $\langle X_{13}Z_{13} \rangle$, $\langle Y_{23}Z_{23} \rangle$, $\langle X_{12} \rangle$, $\langle X_{13} \rangle$, $\langle Y_{12} \rangle$, $\langle Y_{23} \rangle$, $\langle Z_{13} \rangle$, $\langle Z_{23} \rangle$ is*

$$
s_1 \left(\langle X_{12}Y_{12} \rangle, \langle X_{13}Z_{13} \rangle, \langle Y_{23}Z_{23} \rangle, \langle X_{12}X_{13} \rangle, \langle Y_{12}Y_{23} \rangle, \langle Z_{13}Z_{23} \rangle \right) \leq 4.
\tag{45}
$$

where

$$
\begin{aligned}
s_1 (a, b, c, d, e, f) = \max\{\, &(\pm a \pm b \pm c \pm d \pm e \pm f): \\
&\text{the number of minuses is odd}\,\}.
\end{aligned}
\tag{46}
$$

15

Proof. The details are analogous to those of the proof of Lemma 8.1. The polytope in terms of probabilities is defined by 12 equations and 56 inequalities. The 12 equations correspond to the requirement that the 1-marginals of the observable probabilities agree with those of the connections and that the observable probabilities are properly normalized. Expressing the probabilities in terms of the observable and connection expections, the 16 equations become identically true and of the 56 inequalities, 32 turn into those represented by (46) and the remaining 24 correspond to the trivial constraints of the form (41) for the 6 pairs of random variables appearing in (46). □

Lemma 8.5. *If the connection expectations* $\langle X_{12}X_{13}\rangle, \langle Y_{12}Y_{23}\rangle, \langle Z_{13}Z_{23}\rangle$ *are compatible with the observed expectations*

$$\langle X_{12}Y_{12}\rangle, \langle X_{13}Z_{13}\rangle, \langle Y_{23}Z_{23}\rangle, \langle X_{12}\rangle, \langle X_{13}\rangle, \langle Y_{12}\rangle, \langle Y_{23}\rangle, \langle Z_{13}\rangle, \langle Z_{23}\rangle,$$

then, with Δ' defined as in (27),

$$
\begin{aligned}
\Delta' &\geq -\tfrac{1}{2} + \tfrac{1}{2}s_1\left(\langle X_{12}Y_{12}\rangle, \langle X_{13}Z_{13}\rangle, \langle Y_{23}Z_{23}\rangle\right), \\
\Delta' &\geq \tfrac{1}{2}\left(|\langle X_{12}\rangle - \langle X_{13}\rangle| + |\langle Y_{12}\rangle - \langle Y_{23}\rangle| + |\langle Z_{13}\rangle - \langle Z_{23}\rangle|\right), \\
\Delta' &\leq 3 - \left[-\tfrac{1}{2} + \tfrac{1}{2}s_0\left(\langle X_{12}Y_{12}\rangle, \langle X_{13}Z_{13}\rangle, \langle Y_{23}Z_{23}\rangle\right)\right], \\
\Delta' &\leq 3 - \tfrac{1}{2}\left(|\langle X_{12}\rangle + \langle X_{13}\rangle| + |\langle Y_{12}\rangle + \langle Y_{23}\rangle| + |\langle Z_{13}\rangle + \langle Z_{23}\rangle|\right).
\end{aligned}
\tag{47}
$$

Conversely, if these inequalities are satisfied for a given value of Δ', then the connection expectations $\langle X_{12}X_{13}\rangle, \langle Y_{12}Y_{23}\rangle, \langle Z_{13}Z_{23}\rangle$ can always be chosen so that they are compatible with the observable expectations

$$\langle X_{12}Y_{12}\rangle, \langle X_{13}Z_{13}\rangle, \langle Y_{23}Z_{23}\rangle, \langle X_{12}\rangle, \langle X_{13}\rangle, \langle Y_{12}\rangle, \langle Y_{23}\rangle, \langle Z_{13}\rangle, \langle Z_{23}\rangle$$

and yield the given value of Δ' in (27).

Proof. The details are analogous to those of the proof of Lemma (8.2). □

References

Adenier, G. and Khrennikov, A. Yu. (2007). Is the fair sampling assumption supported by EPR experiments? *Journal of Physics B* **40**, pp. 131–141.

Aerts, D., Gabora, L. and Sozzo, S. (2013). Concepts and their dynamics: A quantum theoretical model, *Topics in Cognitive Science* **5**, pp. 737–772.

Aspect, A., Grangier, P. and Roger, G. (1981). Experimental tests of realistic local theories via Bell's theorem, *Physical Review Letters* **47**, pp. 460–463.

Aspect, A., Grangier, P. and Roger, G. (1982). Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of bell's inequalities, *Physical Review Letters* **49**, pp. 91–94.

Badziąg, P., Bengtsson, I., Cabello, A. and Pitowsky, I. (2009). Universality of state-independent violation of correlation inequalities for noncontextual theories, *Physical Review Letters* **103**, p. 050401.

Bell, J. S. (1964). On the Einstein Podolsky Rosen paradox, *Physics* **1**, pp. 195–200.

Cabello, A. (2013). Simple explanation of the quantum violation of a fundamental inequality, *Physical Review Letters* **110**, p. 060402.

Cereceda, J. (2000). Quantum mechanical probabilities and general probabilistic constraints for Einstein–Podolsky–Rosen–Bohm experiments, *Foundations of Physics Letters* **13**, pp. 427–442.

Clauser, J. F., Horne, M. A., Shimony, A. and Holt, R. A. (1969). Proposed experiment to test local hidden-variable theories, *Physical Review Letters* **23**, pp. 880–884.

Dzhafarov, E. N. (2003). Selective influence through conditional independence, *Psychometrika* **68**, 1, pp. 7–25.

Dzhafarov, E. N. and Kujala, J. V. (2010). The joint distribution criterion and the distance tests for selective probabilistic causality, *Frontiers in Psychology* **1**, p. 151, doi:10.3389/fpsyg.2010.00151.

Dzhafarov, E. N. and Kujala, J. V. (2012a). Quantum entanglement and the issue of selective influences in psychology: An overview, in J. R. Busemeyer, F. Dubois, A. Lambert-Mobiliansky and M. Melucci. eds., *Quantum Interaction*, *Lecture Notes in Computer Science*, Vol. 7620, pp. 184–195, Springer.

Dzhafarov, E. N. and Kujala, J. V. (2012b). Selectivity in probabilistic causality: Where psychology runs into quantum physics, *Journal of Mathematical Psychology* **56**, pp. 54–63.

Dzhafarov, E. N. and Kujala, J. V. (2013a). All-possible-couplings approach to measuring probabilistic context, *PLoS ONE* **8**, 5, p. e61712, doi:10.1371/journal.pone.0061712.

Dzhafarov, E. N. and Kujala, J. V. (2013b). Order-distance and other metric-like functions on jointly distributed random variables, *Proceedings of the American Mathematical Society* **141**, pp. 3291–3301.

Dzhafarov, E. N. and Kujala, J. V. (2014a). Contextuality is about identity of random variables. *Physica Scripta* **T163**, p. 014009.

Dzhafarov, E. N. and Kujala, J. V. (2014b). Embedding quantum into classical: contextualization vs conditionalization, *PLoS ONE* **9**, 3, p. e92818.

Dzhafarov, E. N. and Kujala, J. V. (2014c). No-forcing and no-matching theorems for classical probability applied to quantum mechanics, *Foundations of Physics* **44**, pp. 248–265.

Dzhafarov, E. N. and Kujala, J. V. (2014d). On selective influences, marginal selectivity, and Bell/CHSH inequalities, *Topics in Cognitive Science* **6**, 1, pp. 121–128.

Dzhafarov, E. N. and Kujala, J. V. (2014e). A qualified Kolmogorovian account of probabilistic contextuality, *Lecture Notes in Computer Science* **8369**, pp. 201–212.

Dzhafarov, E. N. and Kujala, J. V. (2015). Random variables recorded under mutually exclusive conditions: Contextuality-by-default, in *Advances in Cognitive Neurodynamics IV*, pp. 405–410.

Dzhafarov, E. N. and Kujala, J. V. (forthcoming). Probability, random variables, and selectivity, in W. H. Batchelder, H. Colonius, E. Dzhafarov and J. I. Myung. eds., *The New Handbook of Mathematical Psychology*, to be published by Cambridge University Press.

Dzhafarov, E. N., Kujala, J. V. and Larsson, J.-Å. (2015). Contextuality in three types of quantum-mechanical systems, *Foundations of Physics* **45**, pp. 762–782, doi:10.1007/s10701-015-9882-9.

Fine, A. (1982). Hidden variables, joint probability, and the bell inequalities, *Physical Review Letters* **48**, pp. 291–295.

Khrennikov, A. Yu. (2008a). Bell-Boole inequality: Nonlocality or probabilistic incompatibility of random variables? *Entropy* **10**, pp. 19–32.

Khrennikov, A. Yu. (2008b). EPR-Bohm experiment and Bell's inequality: Quantum physics meets probability theory, *Theoretical and Mathematical Physics* **157**, pp. 1448–1460.

Khrennikov, A. Yu. (2009). *Contextual Approach to Quantum Formalism*, Fundamental Theories of Physics 160, Springer, Dordrecht.

Kirchmair, G., Zähringer, F., Gerritsma, R., Kleinmann, M., Gühne, O., Cabello, A., Blatt, R. and Roos, C. (2009). State-independent experimental test of quantum contextuality, *Nature* **460**, pp. 494–497.

Kochen, S. and Specker, F. (1967). The problem of hidden variables in quantum mechanics, *Journal of Mathematics and Mechanics* **17**, pp. 59–87.

Kujala, J. V. and Dzhafarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors, *Journal of Mathematical Psychology* **52**, pp. 128–144.

Kujala, J. V. and Dzhafarov, E. N. (2015). Proof of a conjecture on contextuality in cyclic systems with binary variables, *arXiv:1503.02181* .

Kujala, J. V., Dzhafarov, E. N. and Larsson, J.-Å. (2015). Necessary and sufficient conditions for maximal contextuality in a broad class of quantum mechanical systems, *arXiv:1412.4724* .

Landau, L. J. (1987). On the violation of Bell's inequality in quantum theory, *Physics Letters A* **120**, pp. 54–56.

Larsson, J.-Å. (2002). A Kochen-Specker inequality, *Europhysics Letters* **58**, pp. 799–805.

Laudisa, F. (1997). Contextualism and nonlocality in the algebra of EPR observables, *Philosophy of Science* **64**, pp. 478–496.

Leggett, A. J. and Garg, A. (1985). Quantum mechanics versus macroscopic realism: is the flux there when nobody looks? *Physical Review Letters* **54**, pp. 857–860.

Masanes, Ll., Acin, A. and Gisin, N. (2006). General properties of nonsignaling theories, *Physical Review A* **73**, p. 012112.

Oas, G., de Barros, J. A. and Carvalhaes, C. (2014). Exploring non-signalling polytopes with negative probability, *Physica Scripta* **T163**, p. 014034.

Specker, E. (1960). Die Logik nicht gleichzeitig entscheidbarer Aussagen, *Dialectica* **14**, pp. 239–246, (English translation by M.P. Seevinck available as arXiv:1103.4537).

Spekkens, R. W. (2008). Negativity and contextuality are equivalent notions of nonclassicality, *Physical Review Letters* **101**, p. 020401.

Suppes, P. and Zanotti, M. (1981). When are probabilistic explanations possible? *Synthese* **48**, pp. 191–199.

Townsend, J. T. and Schweickert, R. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks, *Journal of Mathematical Psychology* **33**, 3, pp. 309 – 327.

Tsirelson, B. S. (1980). Quantum generalizations of Bell's inequality, *Letters in Mathematical Physics* **4**, pp. 93–100.

Weihs, G., Jennewein, T., Simon, C., Weinfurter, H., and Zeilinger, A. (1998). Violation of Bell's inequality under strict Einstein locality conditions, *Physical Review Letters* **81**, pp. 5039–5043.